

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
5 April 2001 (05.04.2001)

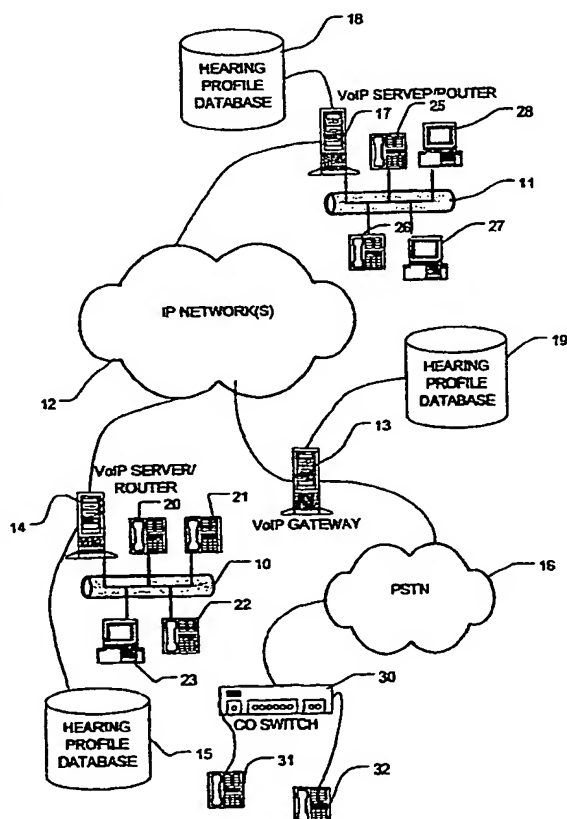
PCT

(10) International Publication Number  
**WO 01/24462 A1**

- (51) International Patent Classification<sup>7</sup>: H04L 12/56 US 09/464,036 (CIP)  
Filed on 15 December 1999 (15.12.1999)
- (21) International Application Number: PCT/US00/26304 US 09/495,255 (CIP)  
Filed on 31 January 2000 (31.01.2000)
- (22) International Filing Date:  
25 September 2000 (25.09.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
09/407,751 28 September 1999 (28.09.1999) US  
09/464,036 15 December 1999 (15.12.1999) US  
09/495,255 31 January 2000 (31.01.2000) US
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:  
US 09/407,751 (CIP)  
Filed on 28 September 1999 (28.09.1999)
- (71) Applicant (for all designated States except US): RX SOUND [US/US]; 801 Welch Road, Palo Alto, CA 94304 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): PLUVINAGE, Vincent [BE/US]; 49 Spencer Lane, Atherton, CA 94027 (US).
- (74) Agent: HAYNES, Mark, A.; Haynes & Beffel LLP, P.O. Box 366, Half Moon Bay, CA 94019 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR DELIVERING CUSTOMIZED VOICE AUDIO DATA ON A PACKET-SWITCHED NETWORK



(57) Abstract: Customized audio for VoIP sessions, or other packet-switched networks for carrying voice, is processed according to a hearing profile of the participants. A processor (13, 14, 17) coupled to a packet-switched communication network (12) processes a flow of packets carrying voice data according to the hearing profiles of the intended recipients. The flow of packets in various embodiments is compliant with a standard Internet Protocol, including, for example, ITU-T Recommendation H.323. The hearing profile data is stored in one embodiment with a registry of users (15, 18, 19). Resources are included in this embodiment to associate data from the registry characterizing a hearing profile of the user with a flow of packets in the packet-switched network for use in customizing the voice data.

WO 01/24462 A1



LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— With international search report.

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**SYSTEM AND METHOD FOR DELIVERING CUSTOMIZED  
VOICE AUDIO DATA ON A PACKET-SWITCHED NETWORK**

**BACKGROUND OF THE INVENTION**

**Field of the Invention**

The present invention relates to the distribution and customization of voice audio data products, like telephone signals, via a packet-switched network such as networks based on the Internet Protocol (IP).

**Description of Related Art**

The hearing profiles of individuals vary in a number of ways. The ability to hear sounds varies with frequency among individuals across the normal audio frequency range. Also, the dynamic range varies among individuals so that levels of an audio stimulus that are perceived as soft sounds and levels of an audio stimulus that are perceived as loud sounds differ from person to person. Standard hearing tests are designed to produce an audiogram that characterizes such factors as frequency, sensitivity and dynamic range in the hearing profiles of individuals. There are also other factors that affect a hearing profile. For example, psycho-acoustic factors concerning the manner in which a person perceives combinations of normal sounds affect the ability to hear in ways that can vary from person to person. Also, environmental factors such as the usual listening environment of a person (library, conference room, concert hall) and the equipment on which the sound is produced (loud speakers, ear phones, telephone hand set) are important. In persons wearing hearing aids or using other assistive hearing devices, the type of aid or device affects the hearing profile. The physiology of an impairment suffered by the individual may also be an important factor in the hearing profile.

The hearing profiles of individuals have been applied in the hearing aid field for customizing and fitting hearing aids for individuals. See, for example, U.S. Patent No. 4,731,850 entitled PROGRAMMABLE DIGITAL HEARING AID SYSTEM, invented by Levitt et al.; and U.S. Patent No. 5,848,171 entitled HEARING AID DEVICE INCORPORATING SIGNAL PROCESSING TECHNIQUES, invented by Stockham, Jr. et al. Thus, techniques for processing sound to offset variations in hearing are well known. However, these techniques are unavailable

to persons not using hearing aids. Furthermore, many persons who could benefit from such processing are not in positions to use hearing aids for a variety of reasons.

Telephone signals are distributed in the public switched telephone networks (PSTNs) in a form intended for persons having hearing profiles within a normal range. Proposals for applying custom processing of telephone signals in the PSTN according to a customer's hearing impairment have been presented. See, for example, U.S. Patent No. 5,388,185 entitled SYSTEM FOR ADAPTIVE PROCESSING OF TELEPHONE VOICE SIGNALS, invented by Terry et al. Such proposals have been limited to circuit-switched telephone networks with limitations inherent in such systems.

Techniques are being developed for the transmission of real time voice over IP based networks, which are sometimes referred to as Voice over Internet Protocol (VoIP) technologies. One leading VoIP technology is based on ITU-T Recommendation H.323, "Packet-Based Multimedia Communications Systems," from the International Telecommunication Union which describes a set of protocols useful for supporting real time audio and video communication over IP networks.

Accordingly, it is desirable to provide systems and methods to apply techniques for the processing of real time audio data according to hearing profiles for the benefit of individuals who are not wearing hearing assistance devices. Also it is desirable to provide tools to simplify the gathering of information needed to develop hearing profiles of individuals, and to apply the gathered information in the distribution and customization of real time voice signals.

### SUMMARY OF THE INVENTION

The present invention provides customized audio for VoIP sessions, or other packet-switched networks for carrying voice, and systems and methods for producing the customized audio data for such sessions, processed according to a hearing profile of the participants. According to the present invention, the quality of VoIP telephone calls and other voice sessions on packet-switched networks is improved.

Thus, the present invention provides a processor coupled to a packet-switched communication network which processes a flow of packets carrying voice data according to the hearing profiles of the intended recipients. The flow of packets in various embodiments is

compliant with a standard Internet Protocol. In one embodiment, the flow of packets is compliant with ITU-T Recommendation H.323. The hearing profile data is stored in one embodiment with a registry of users. Resources are included in this embodiment to associate data from the registry characterizing a hearing profile of the user with a flow of packets in the packet-switched network for use and customizing the voice data.

The invention includes a method for communication of voice signals in a packet-switched network that includes detecting a flow of packets carrying the voice signals on the network to a recipient, utilizing data within the packets. In addition, the method includes obtaining a hearing profile of the recipient from memory on at least one of an intermediate station or an end station in the network. Finally, the flow of packets is processed according to the hearing profile in a processor on one of an intermediate station and an end station in the network. A typical session on a packet-switched network carrying voice includes a first flow from a first participant in the call to the second adjustment in the call, and a second flow from the second participant to the first. The processing of the first and second flows can be positioned in different locales in the network. For example, a flow to a particular participant may be processed in a processor within the local area network of the participant. Data sent from the other participant in the session is delivered to the processor within the local area network of the participant via the Internet for processing prior to delivery. In this manner, the sending station need not have the hearing profile data available for the intended recipients. In alternative systems, the processing of hearing profile data is executing a single server within the network which stores the hearing profiles of the recipients of the data.

According to another embodiment of the invention, the method includes establishing a registry of users in a processor coupled to the network. The registry includes parameters indicating whether processing according to a hearing profile of a user is enabled for that user, and other parameters necessary for establishing the session and providing a hearing profile for use within the session. Accordingly, a processor detects a flow of packets carrying voice signals on the network to a recipient, the packets being identifiable by parameters within the packets as members of the flow. Next, the processor accesses the registry in response to the flow and if the recipient is a registered user, a hearing profile is obtained using parameters available from the registry from memory in the network. Finally, the flow of packets is processed according to a hearing profile.

In another embodiment, the invention includes maintaining the registry in a gate keeper processor on the network, wherein the gate keeper processor performs functions compliant with ITU-T Recommendation H.323.

Accordingly, the present invention provides for the customization of voice sessions in packet-switched networks. A customization of voice sessions requires the management of user registries and of processing resources necessary for matching individual voice sessions with data relevant to the participants in the sessions. The present invention, by taking advantage of the power and flexibility of packet-switched networks such as the Internet, makes the delivery of voice sessions customized according to individual hearing profiles practical and cost-effective.

Other aspects and advantages of the present invention can be seen on review of the drawings, the detailed description, and the claims which follow.

### BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 illustrates a network serving voice over the Internet protocol with customized processing according to the hearing profile of users according to the present invention.

Fig. 2 is a simplified flow chart illustrating the processing at a locale in the network of a voice over the Internet Protocol session with user hearing profiles.

Fig. 3 is a simplified block diagram of an H.323 standard network enhanced with hearing profile processing according to the present invention.

Fig. 4 illustrates one example entry in a registry of hearing profiles for the system of Fig. 3.

Fig. 5 is a simplified flow chart of a process for establishing an H.323 call with customized processing according to the present invention.

Fig. 6 is a diagram of one example of a hearing profile for use in producing customized audio for VoIP sessions.

Fig. 7 is a diagram of another example of a hearing profile for use in producing customized audio for VoIP sessions.

Fig. 8 is a diagram of yet another example of a hearing profile for use in producing customized audio for VoIP sessions.

### DETAILED DESCRIPTION

A detailed description of the various embodiments of the present invention is provided with reference to Figs. 1 through 8.

Fig. 1 provides an illustration of one example network supporting VoIP telephone calls with audio data processed according to the hearing profiles of the participants in the call. The example network includes local area network 10 and local area network 11 which are coupled to one another through an Internet protocol network 12. Also, a gateway server 13 is coupled to the Internet protocol network 12 and to the public switched telephone network PSTN 16. Local area network 10 includes VoIP server/router 14 which couples the local area network 10 to the Internet protocol network 12. Local area network 11 includes a VoIP server/router 17 which couples the local area network 11 to the Internet protocol network 12.

End stations illustrated in the example of Fig. 1 include the telephones 20, 21, 22 coupled to the local area network 10, and the voice enabled personal computer 23 coupled to the local area network 10. Also, the stations coupled to the local area network 11 include telephones 25 and 26, and voice or multimedia enabled personal computers 27 and 28. In the public switched telephone network 16, a central office switch 30 connects end station telephones 31 and 32 into the network.

The intermediate station servers 14 and 17 are coupled to hearing profile databases 15 and 18 to support processing of VoIP sessions according to the hearing profiles of the participants in the call. Thus for example, when a call is established between end station 28 on local area network 11, and end station 22 on local area network 10, the VoIP server/router 17 associates the hearing profile of the recipient using end station 28 with the incoming packet flow of the session, and processes the audio data using the hearing profile. Likewise, the VoIP server/router 14 coupled to the local area network 10 associates the hearing profile of the recipient using the telephone 22 with the incoming packet flow of the session, and processes the audio data using the hearing profile as it is delivered to the telephone 22. In a similar manner, when a VoIP session is established between an end station, such as end station 25 on local area network 11, and a telephone coupled to the public switched telephone network 16, such as telephone 32, the VoIP gateway 13 associates the hearing profile of the recipient at the telephone 32 with incoming packets in the VoIP session, and processes the data using the hearing profile. The processed data is then re-formatted for the circuit switched public switched telephone network for delivery to the recipient at the telephone 32.

The hearing profile databases 15, 18, 19 illustrated in the example of Fig. 1 maintain registries of users of the VoIP service. The registries contain the necessary information to support the processing of the packet flows in the VoIP sessions utilizing the hearing profiles of users having entries in the databases.

5       The establishment and maintenance of the hearing profile databases in the packet-switched network can be handled using a variety of network architectures, taking full advantage of the packet-switched network flexibility. For example, a central server may be used for acquiring, updating, and delivering hearing profiles for use in the VoIP sessions supported by the database. Alternatively, distributed servers, each maintaining hearing profile data for a limited  
10       number of users could be utilized which process only incoming packet flows. In yet another alternative, the hearing profiles may be stored locally on the end stations, and no intermediate station need be involved in the processing of the packet flows according to the hearing profiles. A combination of locales in the network may be used to support a particular session, according to the capability of the stations involved in the session.

15       A basic process executed at a locale in the packet-switched network according to the present invention for packet flows in a VoIP call is illustrated in Fig. 2. The locale in the network, either an end station or an intermediate station, detects a VoIP session by filtering the packet flow to detect headers or other control information tagging the flow as a VoIP session (step 50). Upon detection of a VoIP session, a profile of a recipient at the destination end station  
20       is retrieved from memory (step 51). Next, the packet flow is processed according to the hearing profile of the recipient (step 52). Finally, the processed packets are forwarded to the destination end station (step 53).

As mentioned above, one popular suite of protocols developed for packet based multimedia communication systems such as Internet protocol based network systems used for  
25       VoIP sessions is known as ITU-T Recommendation H.323. Recommendation H.323 is incorporated by reference as if fully set forth herein. Fig. 3 provides a simplified illustration of an H.323 network enhanced according to the present invention. In the network of Fig. 3, an H.323 gatekeeper process 100 is coupled to the Internet 101. H.323 compliant end stations 102 and 103 are also coupled to the Internet 101. The H.323 compliant gateway 104 is coupled to  
30       end stations, such as end station 105 which support other protocol suites for the VoIP session.



The gate keeper process 100 includes an H.323 registry 106 which is enhanced according to the present invention with a hearing profile database. Thus, for example, entries in the H.323 registry 106 have a format such as shown in Fig. 4. The format includes a user identifier 120, and H.323 call parameters 121 which are developed according to the standard. A pointer to a hearing profile, or a hearing profile itself is stored in the field 122. The hearing profile processing capability for the user is stored as a parameter 123. A few examples of the hearing profile pointer 122 include a URL, a memory address in local memory on the gate keeper, or a memory address in a particular storage device having another format. The hearing profile processing capability parameter 123 indicates where the processing is available for applying the hearing profile. Thus, parameter 124 indicates that digital signal processing resources are available locally on an end station of the user (e.g., end station 102), parameter 125 indicates that digital signal processing resources are available on a gateway 104 through which the user accesses the network, and parameter 126 indicates that digital signal processing resources at the gate keeper 100 are to be used for packet flows destined to the identified user.

Fig. 5 provides a simplified flow chart of the processing at the gate keeper 100 of Fig. 4 enhanced according to present invention. During call establishment, the gate keeper detects an H.323 call set up session (step 150). The hearing profile processing capability for the identified end stations is determined using the information in the registry (step 151). Next, the hearing profile or hearing profiles involved in the session are forwarded to the processors that are associated with the registered users. Alternatively, the hearing profile is simply retrieved from the registry for use at the gate keeper processor itself (step 152). The H.323 call set up procedures establish call paths for each direction, taking into account the locale in the network for processing of the audio signals with the hearing profiles of the users (step 153). Finally, the call set up is completed (step 154).

Figs. 6, 7 and 8 illustrate alternative formats for the hearing profiles to be stored in the registry 106, or databases 15, 17, 19 in the examples shown in the Figures. The hearing profile shown in Fig. 6 includes a customer ID 250, coefficients of transform equations 251, and an identifier of the playback device type and listening environment 252 in which the audio data product is to be played. In Fig. 7, the hearing profile includes a customer identifier 260, the audiogram for customer 261, listening condition data 262, and psycho-acoustic parameters 263 which relate to the listening characteristics of the customer. The hearing profile shown in Fig. 8

includes a customer identifier 270, and software 271 along with code data structures 272 that in combination provide executable transform code for producing the customized audio data product from the selected audio data product. The three examples of hearing profile formats shown in Figs. 6-8 are representative of a large number of hearing profile formats that could be utilized depending on the type of transform processes being executed, the type of audio products being delivered, and other factors related to the architecture of the system.

The system in various embodiments provides input tools, such as a graphical user interface usable, for example, at a kiosk or accessible via a network, for accepting input data concerning a hearing profile of a customer, either supported by a specialized hearing test device or by other resources for use in the test. The input tools are used for creating a machine readable hearing profile, which is stored in a registry of hearing profiles for use in the production of customized audio data products. Alternatively, the hearing profile is stored on a portable data storage medium for use by the customer, as described above. The related applications incorporated by reference above, include detailed descriptions of technologies for conducting hearing tests and for gathering hearing profile data.

Also, in various embodiments of the invention, the system provides tools for accepting input data characterizing the feedback from the customer. Processing resources are provided for modifying the machine readable hearing profile in response to the feedback data. The modifying of the machine readable hearing profile in one aspect of the invention comprises applying an optimization modeling process to the hearing profile, by which the profiles are improved over time using feedback arising from a variety of customized audio data products and from a variety of listening conditions. The related applications, incorporated by reference above, include detailed descriptions of technologies for gathering feedback and optimizing hearing profile data.

While the present invention is disclosed by reference to the preferred embodiments and examples detailed above, it is to be understood that these examples are intended in an illustrative rather than in a limiting sense. It is contemplated that modifications and combinations will readily occur to those skilled in the art, which modifications and combinations will be within the spirit of the invention and the scope of the appended claims.

CLAIMS

What is claimed is:

1           1.     For a packet-switched communication network, a station comprising:  
2           a processor coupled to the network which processes a flow of packets carrying voice data  
3           according to hearing profiles of intended recipients.

1           2.     The station of claim 1, wherein the flow of packets comprises packets compliant  
2           with a standard Internet Protocol.

1           3.     The station of claim 1, wherein the flow of packets comprises packets compliant  
2           with ITU-T Recommendation H.323.

1           4.     The station of claim 1, including a registry of users, and including resources to  
2           associate data from the registry characterizing a hearing profile of a user in the registry with the  
3           flow of packets.

1           5.     A method for communication of voice signals in a packet-switched network,  
2           comprising:  
3           detecting a flow of packets carrying the voice signals on the network to a recipient, the  
4           packets identifiable as members of the flow;  
5           obtaining a hearing profile of the recipient from memory on at least one of an  
6           intermediate station and an end station in the network; and  
7           processing the flow of packets according to the hearing profile in a processor on at least  
8           one of an intermediate station and an end station in the network.

1           6.     The method of claim 5, including filtering packets received in the processor to  
2           detect members of the flow.

1           7.     The method of claim 5, wherein the flow originates at an originating end station  
2           in the network, and the memory is on the originating end station.

1           8.     The method of claim 5, wherein the memory is on an intermediate station in the  
2 network.

1           9.     The method of claim 5, wherein the flow is transmitted to a destination end  
2 station in the network, and the memory is on the destination end station.

1           10.    A method for communication of voice signals in a packet-switched network,  
2 comprising:

3                establishing a registry of users in a processor coupled to the network, the registry  
4 including parameters indicating whether processing according to a hearing profile of a user is  
5 enabled for the user;

6                detecting a flow of packets carrying the voice signals on the network to a recipient, the  
7 packets identifiable as members of the flow;

8                accessing the registry in response to the flow, and if the recipient is a registered user,  
9 obtaining a hearing profile of the recipient from memory in the network; and

10               processing the flow of packets according to the hearing profile in a processor in the  
11 network.

1           11.    The method of claim 10, including filtering packets in the processor in the  
2 network to detect members of the flow.

1           12.    The method of claim 10, wherein the flow originates at an originating end station  
2 in the network, and the memory is on the originating end station.

1           13.    The method of claim 10, wherein the memory is on an intermediate station in the  
2 network.

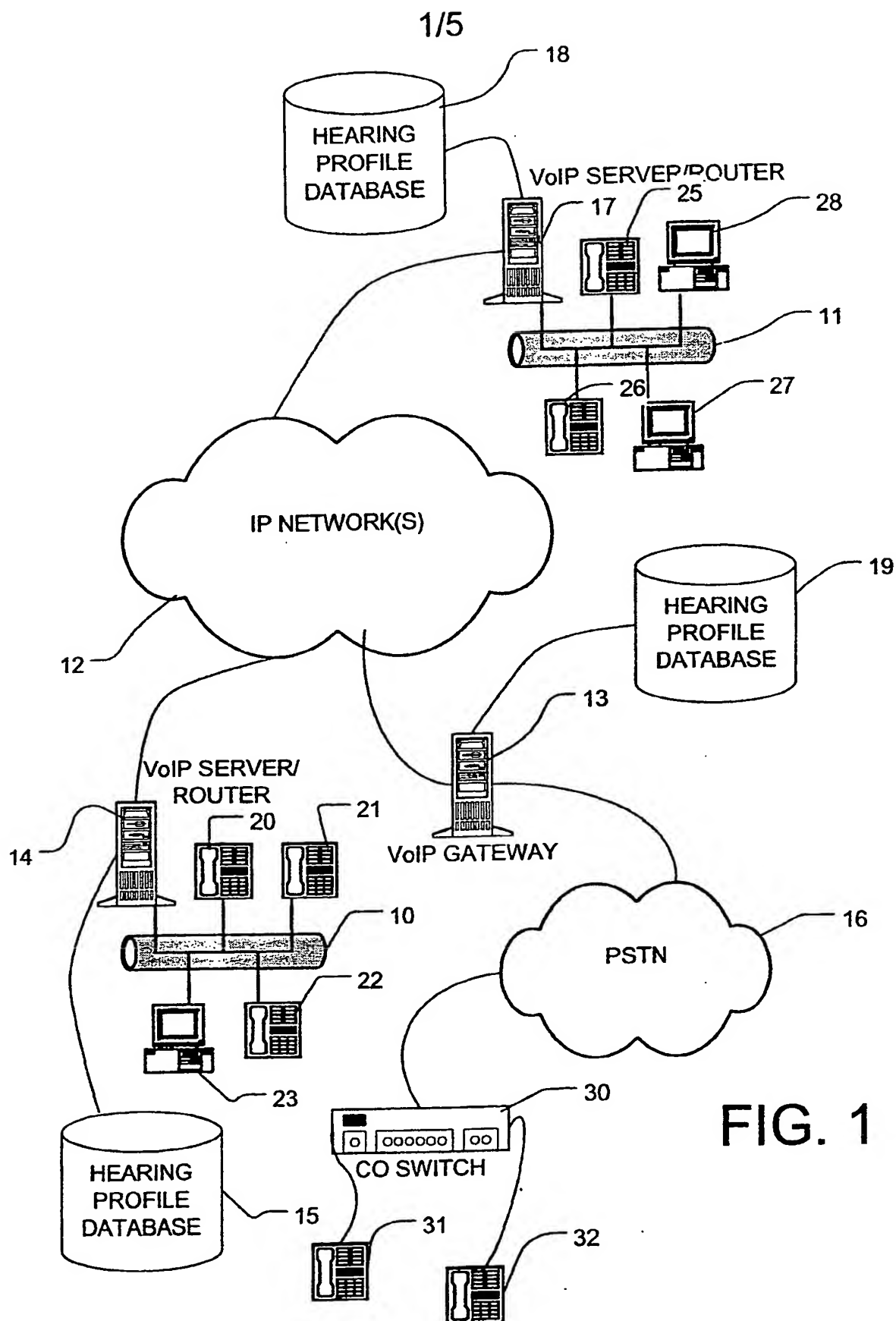
1           14.    The method of claim 10, wherein the flow is transmitted to a destination end  
2 station in the network, and the memory is on the destination end station.

1           15.    The method of claim 10, including maintaining the registry in a gate keeper  
2 processor on the network, and wherein the gate keeper processor performs functions compliant  
3 with ITU-T Recommendation H.323.

1           16.    The method of claim 10, wherein the parameters include a pointer to a hearing  
2 profile for the registered user.

1           17.    The method of claim 10, wherein the processor is on an intermediate station in the  
2 network, and including routing the flow through the intermediate station.

1           18.    The method of claim 10, wherein the processor is on an end station associated  
2 with the recipient, and including transmitting a copy of the hearing profile from the memory in  
3 the network to the end station.



2/5

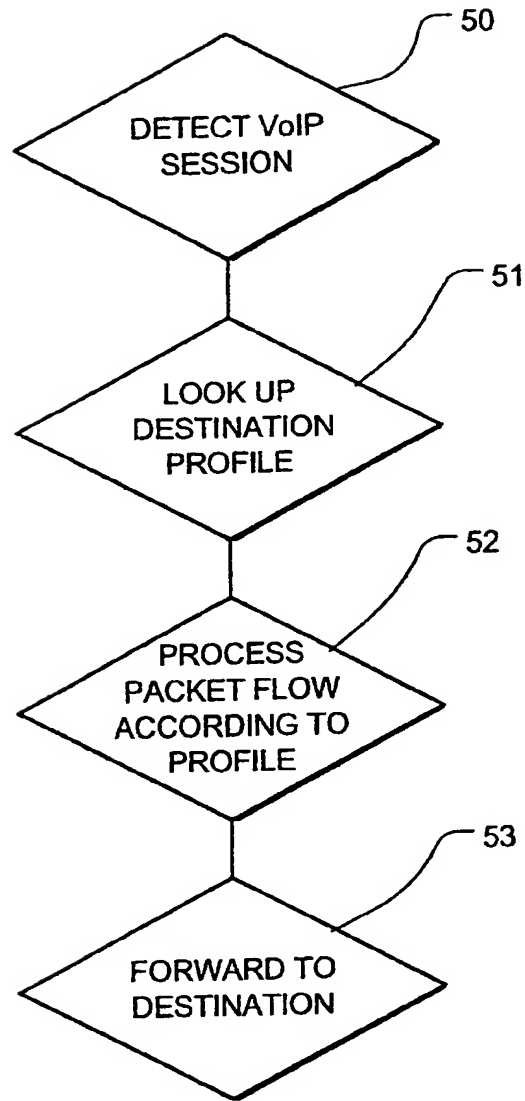
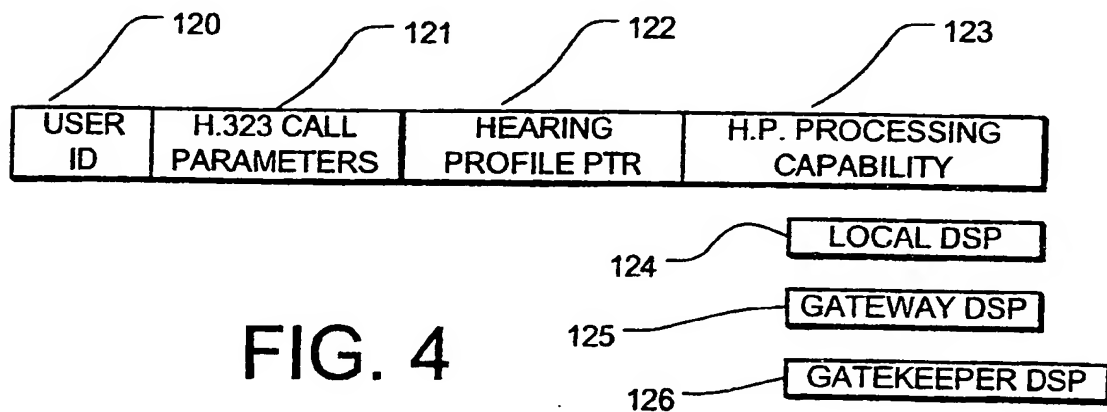
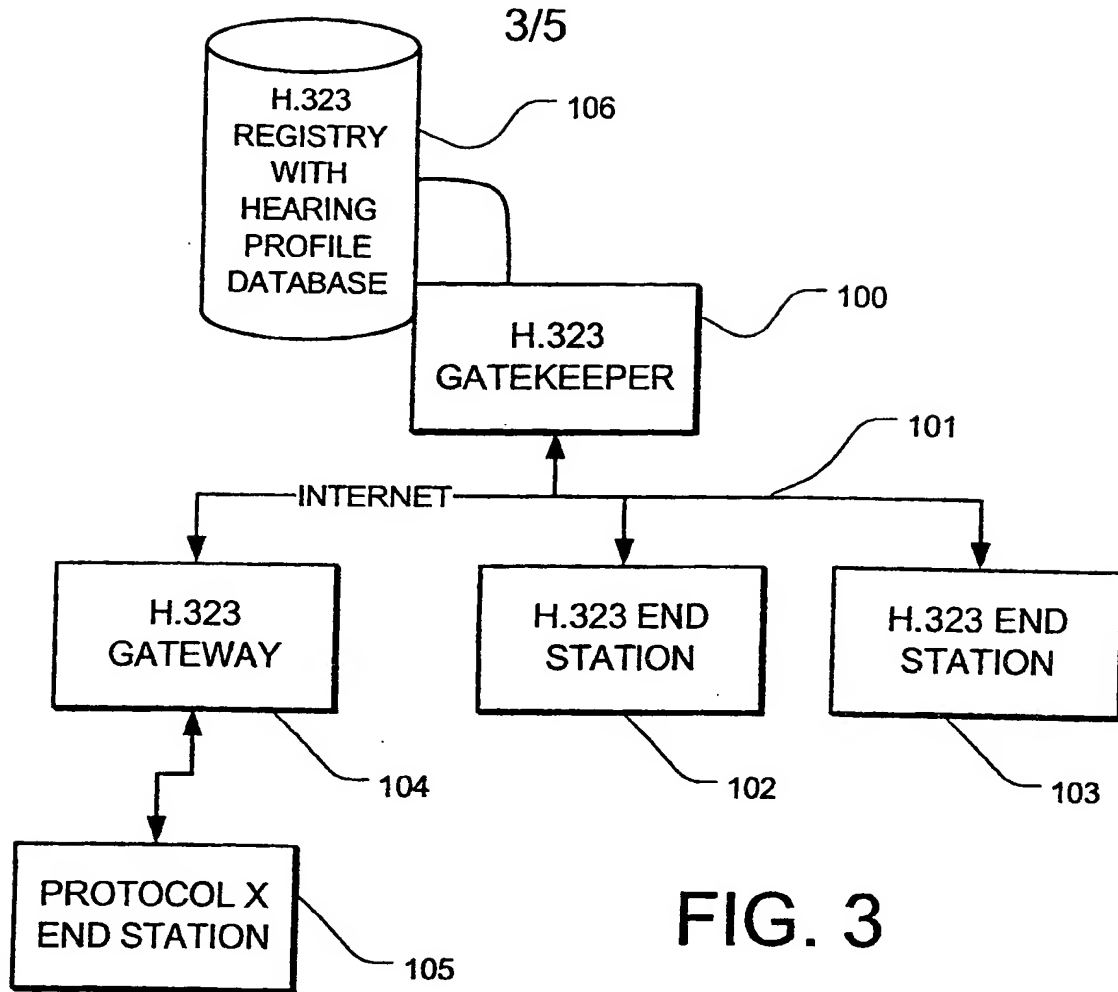


FIG. 2





4/5

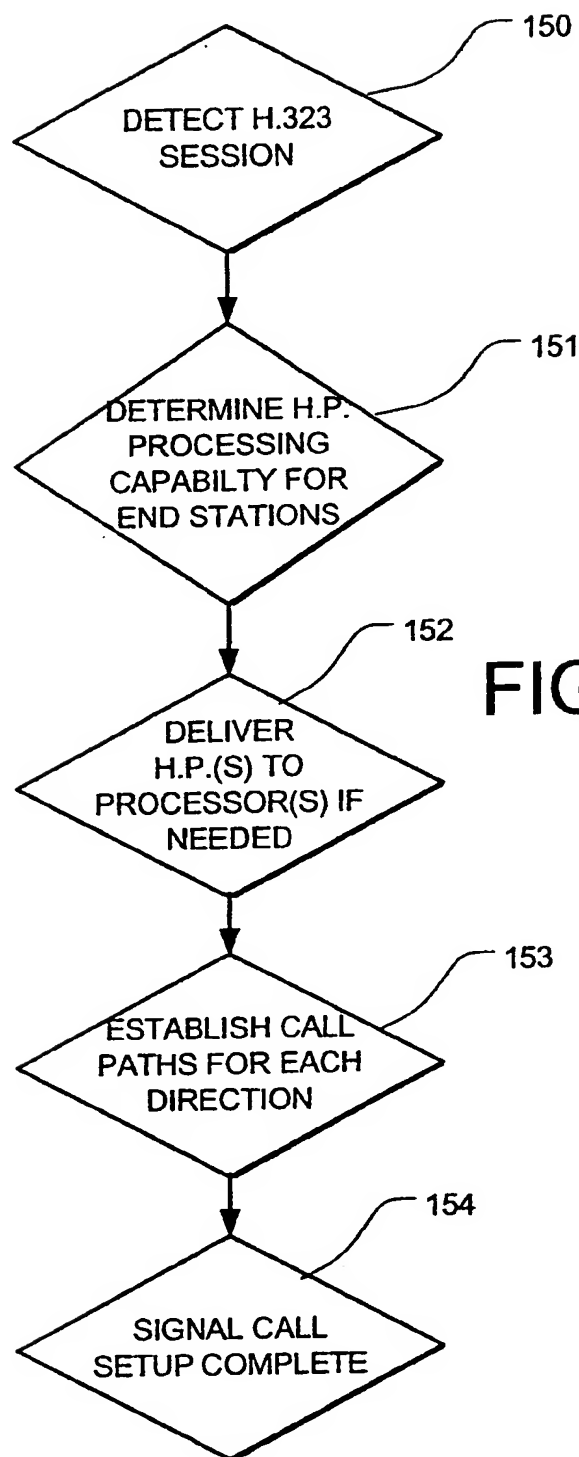


FIG. 5

5/5

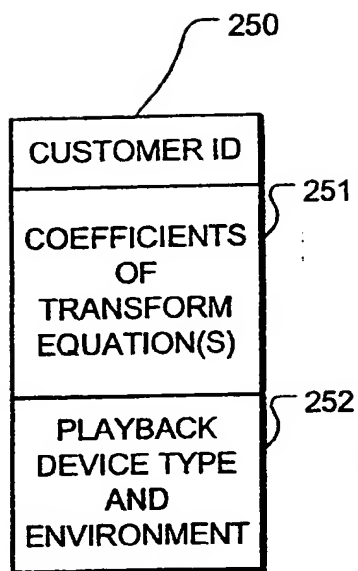


FIG. 6

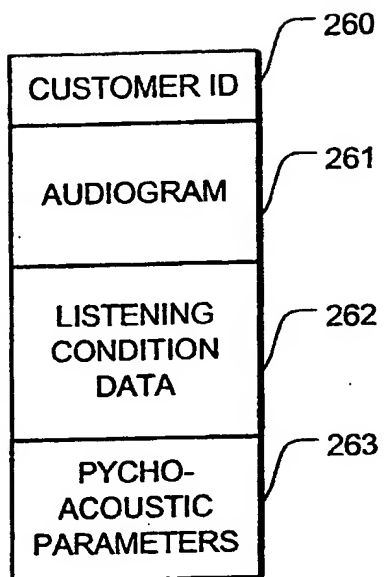


FIG. 7

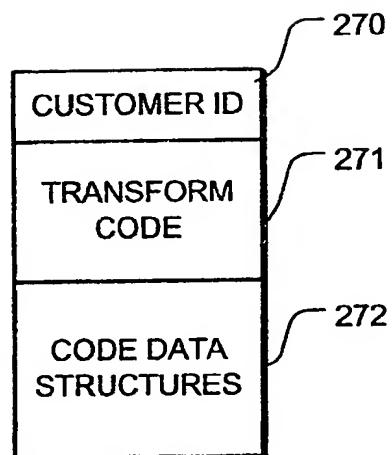


FIG. 8

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/26304

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :H04L 12/56

US CL : 370/395

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 370/395

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,388,185 A (TERRY ET AL) 07 FEBRUARY, 1995, ALL	1-18

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 24 NOVEMBER 2000	Date of mailing of the international search report 19 JAN 2001
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer DOUG OLMS Telephone No. (703) 305-4703